(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 285 449 B1**

# (12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **13.07.94** (51) Int. Cl.5: **G06F 15/20**

(21) Application number: **88302966.2**

(22) Date of filing: **31.03.88**

## (54) Document processing system.

(30) Priority: **31.03.87 JP 78503/87**

(43) Date of publication of application:
**05.10.88 Bulletin 88/40**

(45) Publication of the grant of the patent:
**13.07.94 Bulletin 94/28**

(84) Designated Contracting States:
**DE FR GB**

(56) References cited:

**SYSTEMS-COMPUTERS-CONTROLS, vol. 12, no. 1, January/February 1981, pages 75-83,Silver Spring, Maryland, US; T. TERANISHI et al.: "An alphanumeric text editorwith graphic editing functions"**

**PATENT ABSTRACTS OF JAPAN, vol. 8, no. 192 (P-298)[1629], 4th September 1984;& JP-A-59 79 342 (NIPPON DENSHIN DENWA KOSHA) 08-05-1984**

(73) Proprietor: **KABUSHIKI KAISHA TOSHIBA**
**72, Horikawa-cho**
**Saiwai-ku**
**Kawasaki-shi Kanagawa-ken 210(JP)**

(72) Inventor: **Iwai, Isamu c/o Patent Division**
**Kabushiki Kaisha Toshiba**
**1-1 Shibaura 1-chome**
**Minato-ku Tokyo 105(JP)**
Inventor: **Fukui, Mika c/o Patent Division**
**Kabushiki Kaisha Toshiba**
**1-1 Shibaura 1-chome**
**Minato-ku Tokyo 105(JP)**
Inventor: **Doi, Miwako c/o Patent Division**
**Kabushiki Kaisha Toshiba**
**1-1 Shibaura 1-chome**
**Minato-ku Tokyo 105(JP)**

(74) Representative: **Freed, Arthur Woolf et al**
**MARKS & CLERK,**
**57-60 Lincoln's Inn Fields**
**London WC2A 3LS (GB)**

EP 0 285 449 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

## Description

The present invention relates to a document processing system which can automatically lay out image data such as figures, tables, and the like at appropriate positions in document data.

In a conventional document processing system such as a wordprocessor, when image data such as figures and/or tables is laid out and embedded in a document created by the system, for example, a blank space is inserted in text data constituting a document, and image data of a figure and/or table is attached to the space. With this method, however, each time document data is edited later, the layout operation must be performed.

A function of setting a blank region in document data, e.g., a blank-open function or a region designation function, is provided to a wordprocessor. With this function, the image region is permanently laid out to a specific portion in a specific page. However, in this method, when document data is edited after the image region is laid out, a reference portion of a figure and/or table in text data is considerably separated from the fixed image region in a page, and the correspondence therebetween cannot be easily grasped. Such a case occurs when document data before the reference portion of the figure and/or table is greatly increased or decreased.

In this manner, when a blank space is formed at a given position on a page, if document data is edited thereafter, the position or size of the space in text data must be changed.

In contrast to this, blank lines are inserted in text data to keep an image region in text data. In this case, if document data is edited, the blank space is moved together with the text data, and the reference portion of a figure and/or table in document data can be prevented from being separated from a layout position of the figure and/or table. In this case, however, if these blank lines extend across two pages, the image region also extends across two pages, and the positions of the blank lines must be changed.

When a document is created using a wordprocessor, image data of an existing figure and/or table is often merged in document data. In this case, the size of image data is checked beforehand, and a space corresponding to the image size must be inserted in the document data, resulting in very cumbersome processing. If the size of the image data is changed, the size of the space in the document data must be changed.

In the conventional system as described above, (1) if an image region is permanently set at a position in a given page of document data, the relation between a text and a figure and/or table is often changed by editing of document data, and a cumbersome operation is required to obtain the correct relation. (2) If blank lines are inserted in text data, a set of blank lines for laying out a figure and/or table often extends across two pages by editing of document data, and a cumbersome operation is required to obtain the correct relation. (3) When a space for fetching image data is set in document data, the size of the image data must be checked, and this operation is also cumbersome. (4) When the size of image data is changed, the size of a space in document data must also be changed.

We acknowledge the disclosure in Systems-Computers-Controls, vol. 12, no. 1, January/February 1981, pages 75-83, Silver Spring, Maryland, US; T. Teranishi et al.: "An alphanumeric text editor with graphic editing functions", of a system in which a combination file is created by combining text with special codes which are labels representing images stored elsewhere, so as to organise subsequent merging of text and images: the codes are not intended to be part of the final document, and their complexity makes the system difficult to use. It is also known from Patent Abstracts of Japan, vol. 8, no. 192, (p-192) [1629], 04.09.1984, & JP-A-59 79 342 to merge image and text data, automatically controlling the size of the image.

It is an object of the present invention to provide a document processing system which does not require, to an editor, i.e., an operator, a cumbersome operation when document data or image data is edited, and which can automatically and appropriately lay out an image region in document data.

The document processing system of the present invention is defined in Claim 1.

According to the present invention, the position at which an image is to be inserted in a document is detected based on key information, such as a number or title of a figure and/or table, which forms part of the image data such as a specific figure and/or table, and part also of the text data. Image data to be inserted are specified in accordance with the key information. In the preferred embodiment, when the insertion position and content of the image data are specified, an image region having a size designated by the image data is laid out in document data. In this case, an image data layout operation is performed in accordance with a predetermined layout rule, and layout of an image region extending across two pages or separation of an image reference portion in text data from the corresponding image data is avoided.

Therefore, when document data, i.e., text data and image data, are edited later, the image data can be laid out to an optimal position without requiring any load to an editor. If text data and image data constituting document data are sepa-

rately stored, key information for specifying image data and its insertion position is extracted from the text data. The image data can be specified and the size thereof can be identified based on this information. Thus, the image data can be laid out in document data without extending across two pages of document data. Therefore, a user (operator) need not perform an operation for forming a space corresponding to a region size of image data, and need not care if image data extends across two pages. Since automatic layout operation can be performed, the user's burden can be greatly reduced, and effective document processing can be performed.

This invention can be more fully understood from the following detailed description when taken in conjunction with the accompanying drawings, in which:

Fig. 1 is a block diagram showing a schematic arrangement of a document processing system according to an embodiment of the present invention;

Fig. 2 is a view showing an example of text data input to the system shown in Fig. 1;

Fig. 3 is a view showing a storage content of image title dictionary in the system shown in Fig. 1;

Fig. 4 is a view showing structural analysis of an image title in the system shown in Fig. 1;

Fig. 5 is a view showing an example of image data input to the system shown in Fig. 1;

Fig. 6 is a flow chart showing image title analysis processing in the system shown in Fig. 1; and

Figs. 7A to 7H are views for explaining two different printing images of document data consisting of text data and image data in the system shown in Fig. 1.

An embodiment of the present invention will be described with reference to the accompanying drawings.

Fig. 1 is a block diagram showing a schematic arrangement of a document processing system according to an embodiment of the present invention.

Input section 1 is used for inputting text data constituting a document, image data such as figures and/or tables, and commands for an edit operation. Text data memory 2 stores the text data input through input section 1. Image data memory 3 stores the image data input at input section 1. Text data analyzing section 4 analyzes the text data stored in text data memory 2 in units of, e.g., sentences to detect predetermined key information in the text data, thereby specifying a position of the key information as a position at which a figure and/or table is to be inserted. The key information is information for referring to or indicating a figure and/or table. Note that in this embodiment, a title of a figure or table, i.e., an image title is used as key information. The image title is extracted by image title extracting section 5 with reference to image title dictionary 6. Document structure memory 7 stores the text data analyzed by text data analyzing section 4 and the key information position obtained as a result of analysis. Image data identifying section 8 retrieves image data having the same image title as that extracted from the text data from image data memory 3, and identifies it. Image size identifying section 9 identifies the region size of the identified image data. Open & layout controller 10 performs layout-control of the image data having the identified size at a position corresponding to the key information position stored in document structure memory 7. The layout control is performed in accordance with an image layout rule stored in layout rule memory 11 while opening or developing text data in a page image. Output section 12 outputs the document data in which image data is laid out. Document processing controller 13 controls the operations of these sections.

Text data input from input section 1 is stored in text data memory 2 as original text data and is also sent to text data analyzing section 4 under the control of document processing controller 13. Text data analyzing section 4 analyzes input text data in units of sentences divided by a period or a return code, and extracts a sentence serving as an image title. For example, assuming that text data shown in Fig. 2 is input, matching processing between sentences and image title dictionary 6 shown in Fig. 3, i.e., processing for determining whether sentences match with the contents of image title dictionary 6, is performed. As a result, a sentence [Fig. 2 Relation Between Text and Image Region] in text data shown in Fig. 2 is analyzed into three categories, i.e., an image classification word, an alphanumeric expression, and a noun phrase, as indicated by A in Fig. 4. In some cases, an image title may include an additional word such as "No." in addition to the above-mentioned three categories like the content of image title dictionary shown in Fig. 3. A category pattern consisting of the categories is analyzed in accordance with an image title rule prestored in image title extracting section 5, and this sentence is analyzed to be an image title, as indicated by B in Fig. 4. The value of key information position indicated by, e.g., a sentence number of the sentence, and a structural attribute indicating the image title are stored in document structure memory 7. The image title rule is as follows.

Image Title Rule

[Rule 1]

Condition 1. A sentence is constituted in the order of an image classification word, an alphanumeric expression, and a noun phrase.

[Result] The sentence is an image title.

[Rule 2]

Condition 1. A sentence is constituted in the order of an image classification word, an additional word, an alphanumeric expression, and a noun phrase.

[Result] The sentence is an image title.

[Rule 3]

Condition 1. A sentence is constituted in the order of an additional word, an alphanumeric expression, an image classification word, and a noun phrase.

[Result] The sentence is an image title.

...(others)

Image data identifying section 8 performs analysis of image data shown in Fig. 5 as in text data analyzing section 4. More specifically, section 8 extracts and identifies an image title in the image data. This identification process is performed in the same manner as in image title identification in text data described above using the image title rule described above with reference to image title dictionary 6 having the content shown in Fig. 3. As a result, if image data including an image title of the same category pattern as that stored in document structure memory 7 is detected, the image data is identified as one to be inserted at key information position P. Fig. 6 shows this processing flow. More specifically, image data is read out from image data memory 3 (step 21). A sentence until a period or a carriage return is extracted (step 22). It is checked if the extracted sentence coincides with the content of image title dictionary 6 (step 23). Matching processing (step 24) for determining whether or not the category pattern of the sentence coincides with any of the image title rules, and matching processing (step 25) for determining whether or not the category pattern of the image title extracted from the image data coincides with the category pattern of the image title extracted from the text data, are sequentially performed. As a result of these matching processing operations, if a coincidence is found, the corresponding image data is transferred to image size identifying section 9 (steps 26 and 27).

Image size identifying section 9 identifies the size of the image data. The size of the image data is, for example, a portion indicated by a broken line in Fig. 5, and the size information is stored in image data memory 3 to be paired with the image data. Document processing controller 13 writes the identified image size at key information position P indicated by the sentence number in document structure memory 7.

When the relation between the image title in the text data and the image data, and its image size are extracted, open & layout controller 10 develops the text data and image data in document data of a printing image. The text data is sequentially read out from the beginning of the sentence, and is laid out in the printing image as shown in any one of Figs. 7A to 7H. Note that Figs. 7A to 7H illustrate examples wherein data is developed in two columns per page. The structural attribute of each sentence in document structure memory 7 is then referred. When a Pth sentence is read, it can be detected that the structural attribution of the sentence is an image title. In this case, the size of the image data is checked, and the image layout rule is read out from image layout rule memory 11 to check a layout position. Image layout rules stored in image layout rule memory 11 are as follows.

Image Layout Rule

[Rule 1]

Condition 1. An X-size of an image region exceeds an X-size of a text frame.

Condition 1.1. Key information of image data is located in a lower half in the Y direction of the text frame.

[Result] The image region is arranged at the lowermost position of the frame.

Condition 1.2. Key information of image data is located in an upper half in the Y direction of the text frame.

Condition 1.2.1. A frame has a plurality of columns, and a blank frame is present to the right of the corresponding frame.

[Result] The image region is arranged at the uppermost position of the right blank frame.

Condition 1.2.2. A frame has one or a plurality of columns, and there is no blank frame to the right of the corresponding frame.

[Result] The image region is arranged at the uppermost position of the starting frame in the next page.

Condition 2. An X-size of an image region is equal to or smaller than an X-size of the text frame.

Condition 2.1. Y-size yn of image data is equal to or smaller than size yf of the remain-

ing portion of the corresponding frame.

[Result] The image region is arranged below a sentence of sentence number P-1.

Condition 2.2. Y-size yn of image data is larger than size yf of the remaining portion of the corresponding frame.

Condition 2.2.1. A frame has a plurality of columns, and a blank frame is present to the right of the corresponding frame.

[Result] The image region is arranged at the uppermost position of the right blank frame.

Condition 2.2.2. A frame has one or a plurality of columns, and there is no blank frame to the right of the corresponding frame.

[Result] The image region is arranged at the uppermost position of the starting frame in the next page.

Condition 3. Y-size yn of image data is larger than Y-size yf of a frame or Y-size yp of a page.

[Result] Y- and X-sizes of the image region are reduced to (yf/yn) or (yp/yn), and the image region is arranged.

In the case of Fig. 7A, condition 2.1 of rule 1 is applied, and an image region is laid out to the left one of the two columns. If Y-size yn of image data is larger than remaining size yf of the left frame size, as shown in Fig. 7B, image data cannot be inserted, and condition 2.2.1 of rule 2 is applied, and the P + 1 sentence is read out and developed after the P-1 sentence but prior to the P sentence. Image data corresponding to an image title of sentence number P is developed in a right frame. More specifically, sentences are reordered, so that an image data region does not extend across two frames. These processing operations are performed by open & layout controller 10. After data is developed in a printed image, the obtained data is output to output section 12.

As shown in Fig. 7C, distance py from the upper end of the frame to key information position P is larger than a size half Y-size yf of the frame [py ≥ (yf/2)], condition 1.1. is applied, and an image region is arranged at the lowermost position of the frame, as shown in Fig. 7D.

Opposite to Fig. 7C, as shown in Fig. 7E, when distance py from the upper end of the frame to key information position P is smaller than a size half Y-size yf of the frame [py ≤ (yf/2)], condition 1.2.1. is applied, and an image region is arranged at the uppermost position of the right frame, as shown in Fig. 7F.

When Y-size yn of the image region exceeds Y-size yf of the frame or Y-size yp of a page, as shown in Fig. 7G, condition 3 is applied, and the image size is reduced to [(yf/yn) or (yp/yn) times] so that the image region falls within the frame, and the reduced image region is arranged in the frame, as shown in Fig. 7G.

As described above, text data and image data are separately created, and image data can be merged in text data at a position near the image title described in the text data so that an image data region does not extend across two pages, thus allowing effective document processing.

The present invention is not limited to the above embodiment. In the above embodiment, the same image title as that written in image data is written in text data, and is used as key information to determine the relation by matching processing. However, matching processing may be performed using only an image classification word and an alphanumeric expression as key information. Thus, an image title itself need not be written in text data. For example, "Fig. 2" is extracted from text data "as shown in Fig. 2", and its insertion position can be identified. At this time, the insertion position of image data may be laid out to an identical page in which "Fig. 2" is developed, in accordance with a layout rule.

The size of image data is not limited to a size smaller than one page, but may be constituted by one document consisting of a plurality of pages or a plurality of documents. These image data can be identified from document data using a file name or the like. In this case, the file name of the image data may be inserted as background data in the text data, i.e. part of a control data group corresponding to text data.

Furthermore, the user can alter the image title rule and the image layout rule, and can set new rules in place of the rules described above.

## Claims

1. A document processing system comprising:

input means (1) for inputting text data and image data constituting document data, the image data being laid out in the document data;

memory means (2, 3) for storing text data and image data input through said input means (1);

text analyzing means (4, 5, 6) for identifying a position in the document data at which the image data is to be assigned;

image identifying means (8, 5, 6) for identifying image data stored in said memory means (2, 3);

image size identifying means (9) for identifying an image size of the image data identified by said image identifying means (8, 5, 6);

layout processing means (10, 11) for laying out an image region of the identified image data, at a location corresponding to the identified image layout position in the document

data, in accordance with a predetermined layout rule; and

output means (12) for outputting the document data in which the image region is laid out by said layout processing means (10, 11);

characterised in that the system is adapted to process text data and image data linked to a portion of that text by common key information (A, Fig. 4) which forms part of the text and also forms part of the image, the key information being intended for inclusion in the output document as part of the text and also as part of the associated image, the key information having a predetermined recognizable structure or structures;

the text analyzing means (4, 5, 6) including means for analyzing sections of text (Fig. 2) and for determining whether each text section has a predetermined structure expected of key information, and, if so, for recording its position in the document data;

the image identifying means (8, 5, 6) including means for analyzing sections of the image (Fig. 5) and for determining whether each image section has a predetermined structure expected of key information and, if so, for recording its key information, and including means for comparing the key information of the image with the key information of the text and for recording the intended position of the image in the document data in accordance with the result of that comparison.

2. A system according to Claim 1, in which the key information is an alphanumeric title of an image.

3. A system according to Claim 1 or 2, in which said layout processing means comprises means (10) for performing layout control in accordance with a layout rule for controlling a layout position such that an image region does not extend across two pages.

4. A system according to Claim 1, 2 or 3, in which the layout rule is for determining a layout position near the key information in the text data.

5. A system according to any preceding claim, in which said layout processing means includes means (10) for varying the image size before an image region of the identified image data is laid out in the document data.

6. A system according to any preceding claim, in which said layout processing means includes means (10) for altering the layout rule.

**Patentansprüche**

1. Dokumentverarbeitungssystem mit:

einer Eingabe- bzw. Eingangseinrichtung (1) zum Eingeben von Textdaten und Bilddaten, die Dokumentdaten bilden, wobei die Bilddaten in den Dokumentdaten angeordnet sind;

einer Speichereinrichtung (2, 3) zum Speichern von Textdaten und Bilddaten, die durch die Eingangseinrichtung (1) eingegeben sind;

einer Textanalysiereinrichtung (4, 5, 6) zum Identifizieren einer Position in den Dokumentdaten, der die Bilddaten zuzuweisen sind;

einer Bildidentifizierungseinrichtung (8, 5, 6) zum Identifizieren von Bilddaten, die in der Speichereinrichtung (2, 3) gespeichert sind;

einer Bildgrößenidentifizierungseinrichtung (9) zum Identifizieren einer Bildgröße der Bilddaten, die durch die Bildidentifizierungseinrichtung (8, 5, 6) identifiziert sind;

einer Layout- bzw. Anordnungsverarbeitungseinrichtung (10, 11) zum Anordnen einer Bildregion der identifizierten Bilddaten an einer Stelle, die der identifizierten Bildanordnungsposition in den Dokumentdaten entspricht, gemäß einer vorbestimmten Anordnungsregel; und

einer Ausgabe- bzw. Ausgangseinrichtung (12) zum Ausgeben der Dokumentdaten, in denen die Bildregion durch die Anordnungsverarbeitungseinrichtung (10, 11) angeordnet ist;

dadurch gekennzeichnet, daß

das System angepaßt ist, um Textdaten und Bilddaten, die mit einem Teilbereich dieses Textes durch eine gemeinsame Schlüsselinformation (A, Fig. 4) verbunden sind, die einen Teil des Textes bildet und auch einen Teil des Bildes bildet, wobei die Schlüsselinformation für die Aufnahme in das Ausgangsdokument als Teil des Textes und auch als Teil des verknüpften Bildes vorgesehen ist, wobei die Schlüsselinformation eine vorbestimmte, erkennbare Struktur oder Strukturen hat;

die Textanalysiereinrichtung (4, 5, 6) eine Einrichtung hat zum Analysieren von Textabschnitten (Fig. 2) und zum Bestimmen, ob jeder Textabschnitt eine vorbestimmte Struktur hat, die von einer Schlüsselinformation erwartet wird, und, falls dies so ist, zum Aufzeichnen seiner Position in den Dokumentdaten;

die Bildidentifizierungseinrichtung (8, 5, 6) eine Einrichtung beinhaltet zum Analysieren von Bildabschnitten (Fig. 5) und zum Bestimmen, ob jeder Bildabschnitt eine vorbestimmte Struktur hat, die von einer Schlüsselinformation erwartet wird, und, falls dies so ist, zum Aufzeichnen seiner Schlüsselinformation, und eine Einrichtung beinhaltet zum Vergleichen der

Schlüsselinformation des Bildes mit der Schlüsselinformation des Textes und zum Aufzeichnen der vorgesehenen Position des Bildes in den Dokumentdaten gemäß dem Ergebnis dieses Vergleiches.

2. System nach Anspruch 1, bei dem die Schlüsselinformation ein alphanumerischer Titel eines Bildes ist.

3. System nach Anspruch 1 oder 2, bei dem die Anordnungsverarbeitungseinrichtung eine Einrichtung (10) umfaßt zur Durchführung der Anordnungssteuerung gemäß einer Anordnungsregel zum Steuern einer Anordnungsposition, so daß eine Bildregion sich nicht über zwei Seiten erstreckt.

4. System nach Anspruch 1, 2 oder 3, bei dem die Anordnungsregel zum Bestimmen einer Anordnungsposition in der Nahe der Schlüsselinformation in den Textdaten dient.

5. System gemäß irgendeinem vorhergehenden Anspruch, bei dem die Anordnungsverarbeitungseinrichtung eine Einrichtung (10) beinhaltet zum Verändern der Bildgröße, bevor eine Bildregion der identifizierten Bilddaten in den Dokumentdaten angeordnet ist.

6. System gemäß irgendeinem vorhergehenden Anspruch, bei dem die Anordnungsverarbeitungseinrichtung eine Einrichtung (10) zum Ändern der Anordnungsregel beinhaltet.

**Revendications**

1. Système de traitement de document comportant:

des moyens d'entrée (1) pour introduire des données de texte et des données d'image constituant des données de document, les données d'image étant disposées dans les données de document ;

des moyens de mémoire (2, 3) pour mémoriser des données de texte et des données d'image introduites par l'intermédiaire desdits moyens d'entrée (1) ;

des moyens d'analyse de texte (4, 5, 6) pour identifier une position dans les données de document à laquelle les données d'image doivent être affectées ;

des moyens d'identification d'image (8, 5, 6) pour identifier des données d'image mémorisées dans lesdits moyens de mémoire (2, 3) ;

des moyens d'identification de dimension d'image (9) pour identifier une dimension d'image des données d'image identifiées par lesdits moyens d'identification d'image (8, 5, 6) ;

des moyens de traitement de disposition (10, 11) pour disposer une région d'image des données d'image identifiées, en un emplacement correspondant à l'endroit de disposition d'image identifié dans les données de document, conformément à une règle de disposition prédéterminée ; et

des moyens de sortie (12) pour délivrer les données de document dans lesquelles la région d'image est disposée par lesdits moyens de traitement de disposition (10, 11) ;

caractérisé en ce que le système est adapté pour traiter des données de texte et des données d'image liées à une partie de ce texte par une information clé commune (A, Figure 4) qui constitue une partie du texte et constitue également une partie de l'image, l'information clé étant prévue pour une incorporation dans le document de sortie en tant que partie du texte et également en tant que partie de l'image associée, l'information clé possédant une ou des structures reconnaissables prédéterminées ;

les moyens d'analyse de texte (4, 5, 6) comprenant des moyens pour analyser des parties de texte (Figure 2) et pour déterminer si chaque section de texte possède une structure prédéterminée attendue de l'information clé, et, s'il en est ainsi, pour enregistrer sa position dans les données de document ;

les moyens d'identification d'image (8, 5, 6) comprenant des moyens pour analyser des parties de l'image (Figure 5) et pour déterminer si chaque partie d'image possède une structure prédéterminée attendue de l'information clé et, s'il en est ainsi, pour enregistrer son information clé, et comprenant des moyens pour comparer l'information clé de l'image à l'information clé du texte et pour enregistrer la position prévue de l'image dans les données de document conformément au résultat de cette comparaison.

2. Système selon la revendication 1, dans lequel l'information clé est un titre alphanumérique d'une image.

3. Système selon la revendication 1 ou 2, dans lequel lesdits moyens de traitement de disposition comprennent des moyens (10) pour effectuer une commande de disposition conformément à une règle de disposition pour commander un emplacement de disposition de telle sorte qu'une région d'image ne s'étende pas sur deux pages.

4. Système selon la revendication 1, 2, ou 3, dans lequel la règle de disposition a pour objet de déterminer un emplacement de disposition près de l'information clé dans les données de texte.

5. Système selon l'une quelconque des revendications précédentes, dans lequel lesdits moyens de traitement de disposition comprennent des moyens (10) pour modifier la dimension d'image avant qu'une région d'image des données d'image identifiées soit disposée dans les données du document.

6. Système selon l'une quelconque des revendications précédentes, dans lequel lesdits moyens de traitement de disposition comprennent des moyens (10) pour modifier la règle de disposition.

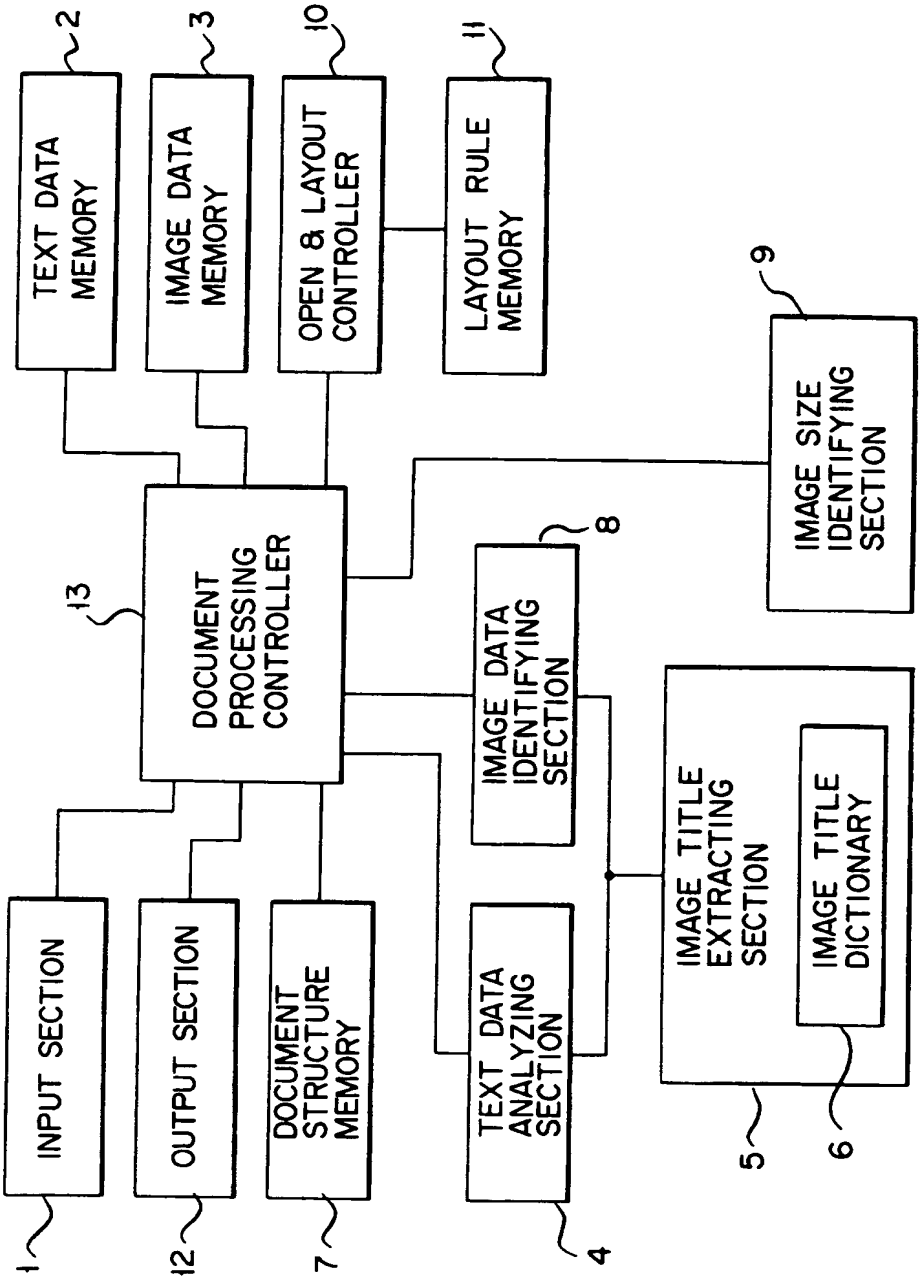
TEXT DATA MEMORY
2
IMAGE DATA MEMORY
3
OPEN & LAYOUT CONTROLLER
10
LAYOUT RULE MEMORY
11
IMAGE SIZE IDENTIFYING SECTION
9
DOCUMENT PROCESSING CONTROLLER
13
IMAGE DATA IDENTIFYING SECTION
8
IMAGE TITLE EXTRACTING SECTION
IMAGE TITLE DICTIONARY
5
6
INPUT SECTION
1
OUTPUT SECTION
12
DOCUMENT STRUCTURE MEMORY
7
TEXT DATA ANALYZING SECTION
4

FIG. 1

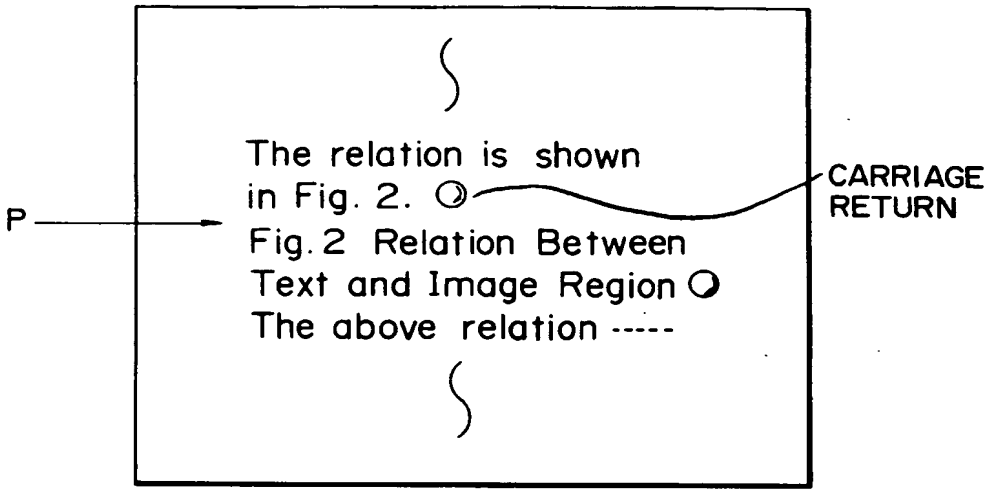


FIG. 2

| CATEGORY | ITEM |
|---|---|
| IMAGE CLASSIFICATION WORDS | Fig.<br>Table<br>Figure<br>⋮ |
| ALPHANUMERIC EXPRESSIONS | 1,2,3,-------<br>①,②,③,------<br>I, II, III,------<br>A,B,C,------<br>⋮ |
| ADDITIONAL WORDS | No.<br>#<br>⋮ |

FIG. 3

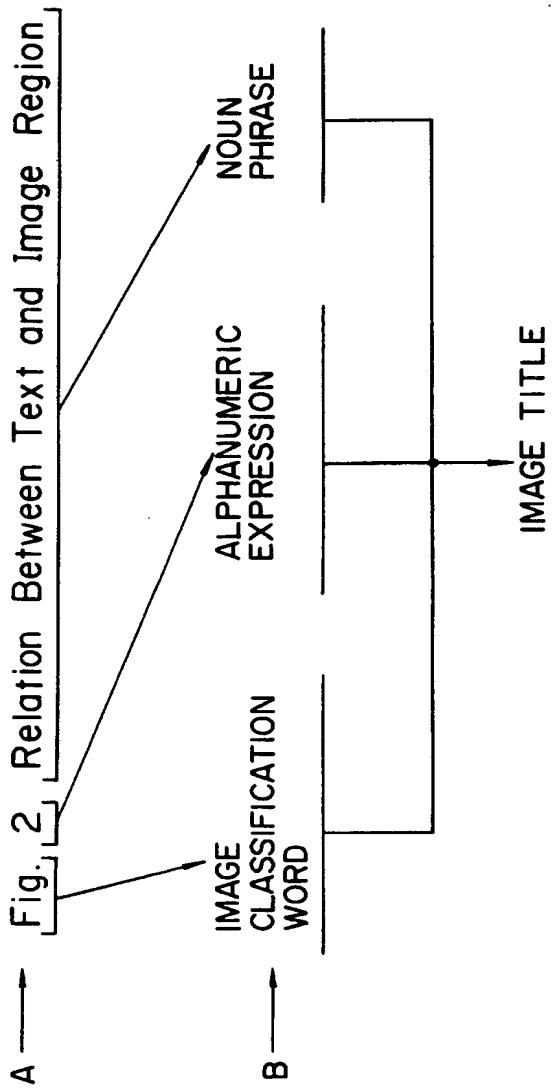
Fig. 2 Relation Between Text and Image Region
A
B
IMAGE CLASSIFICATION WORD
ALPHANUMERIC EXPRESSION
NOUN PHRASE
IMAGE TITLE

FIG. 4

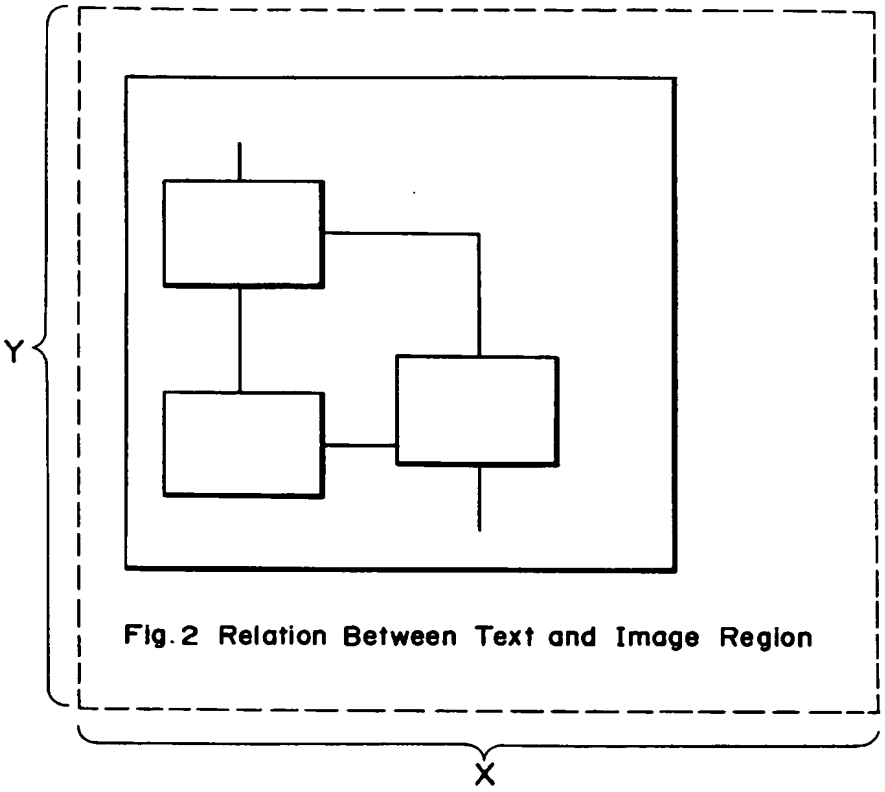


Fig. 2 Relation Between Text and Image Region

FIG. 5

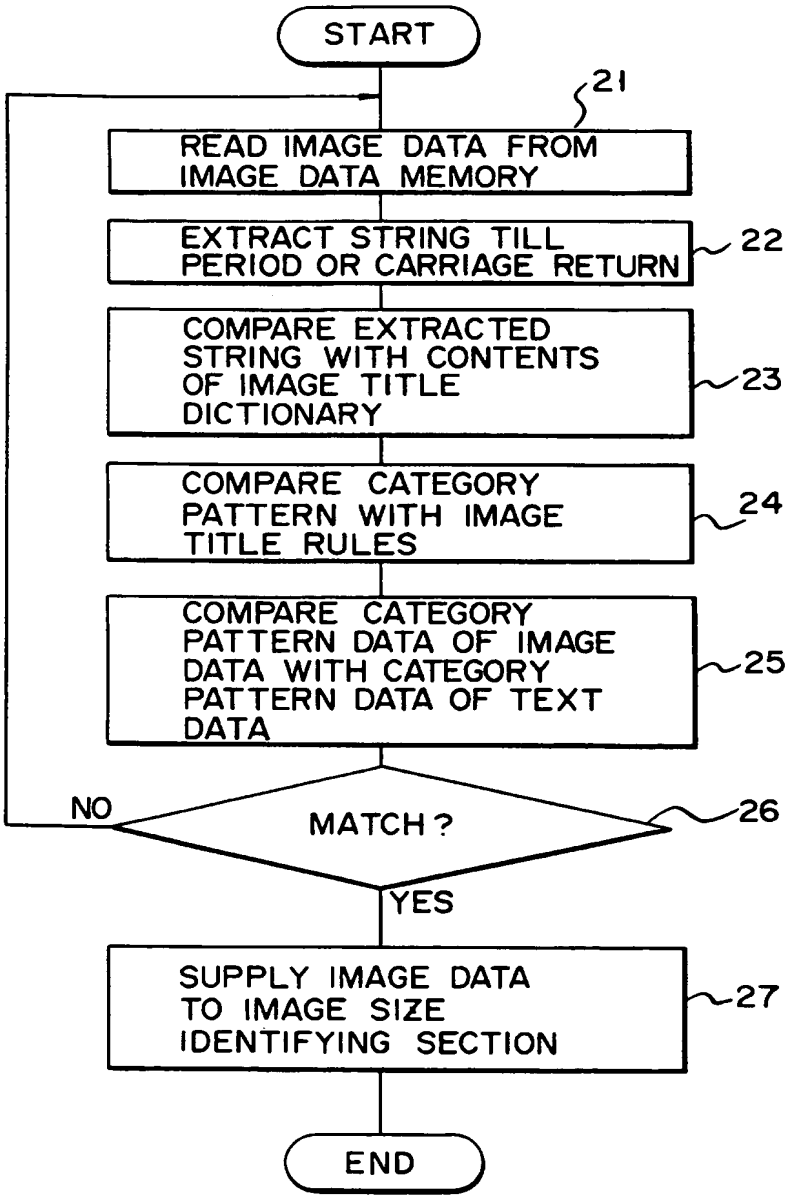
START
21
READ IMAGE DATA FROM IMAGE DATA MEMORY
EXTRACT STRING TILL PERIOD OR CARRIAGE RETURN
22
COMPARE EXTRACTED STRING WITH CONTENTS OF IMAGE TITLE DICTIONARY
23
COMPARE CATEGORY PATTERN WITH IMAGE TITLE RULES
24
COMPARE CATEGORY PATTERN DATA OF IMAGE DATA WITH CATEGORY PATTERN DATA OF TEXT DATA
25
NO
MATCH ?
26
YES
SUPPLY IMAGE DATA TO IMAGE SIZE IDENTIFYING SECTION
27
END

FIG. 6

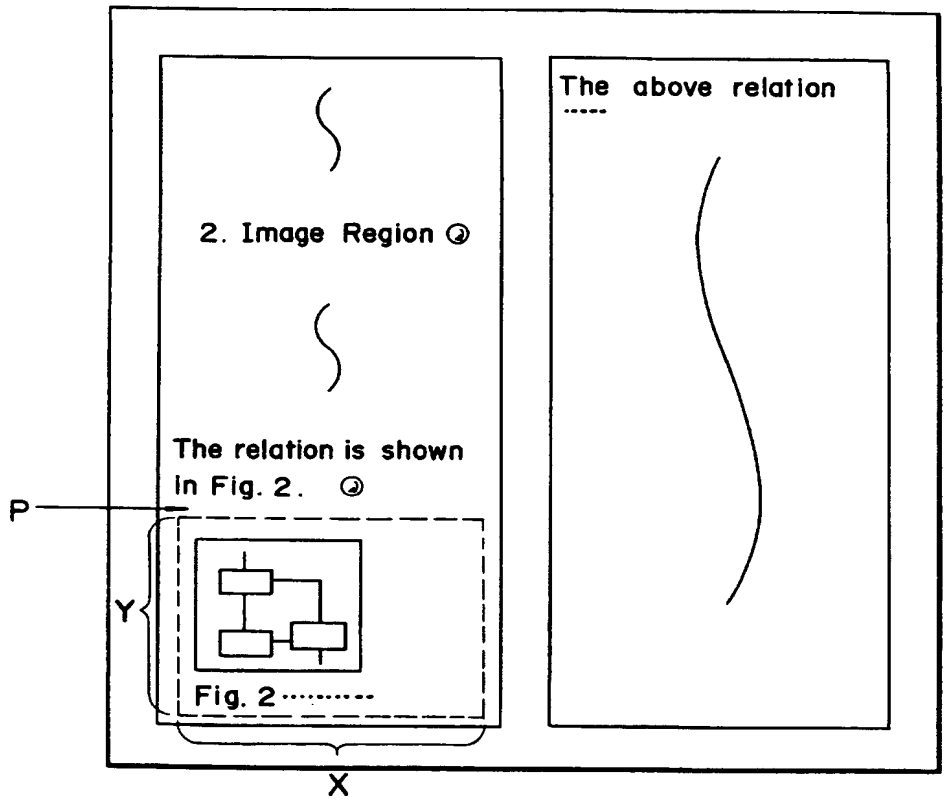
2. Image Region ②
The relation is shown in Fig. 2. ②
P
Y
Fig. 2
X
The above relation

FIG. 7A

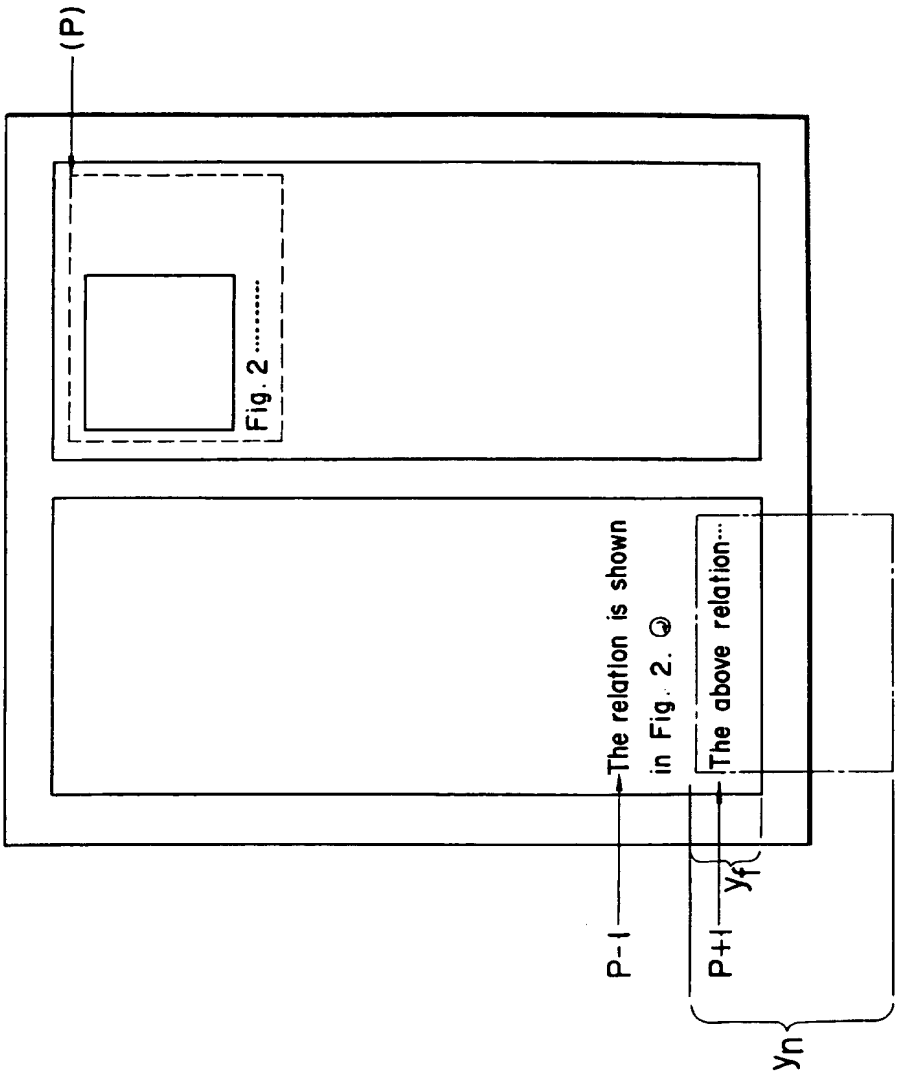
(P)
Fig. 2
The relation is shown in Fig. 2. ②
The above relation···
P-1
P+1
yf
yn

FIG. 7B

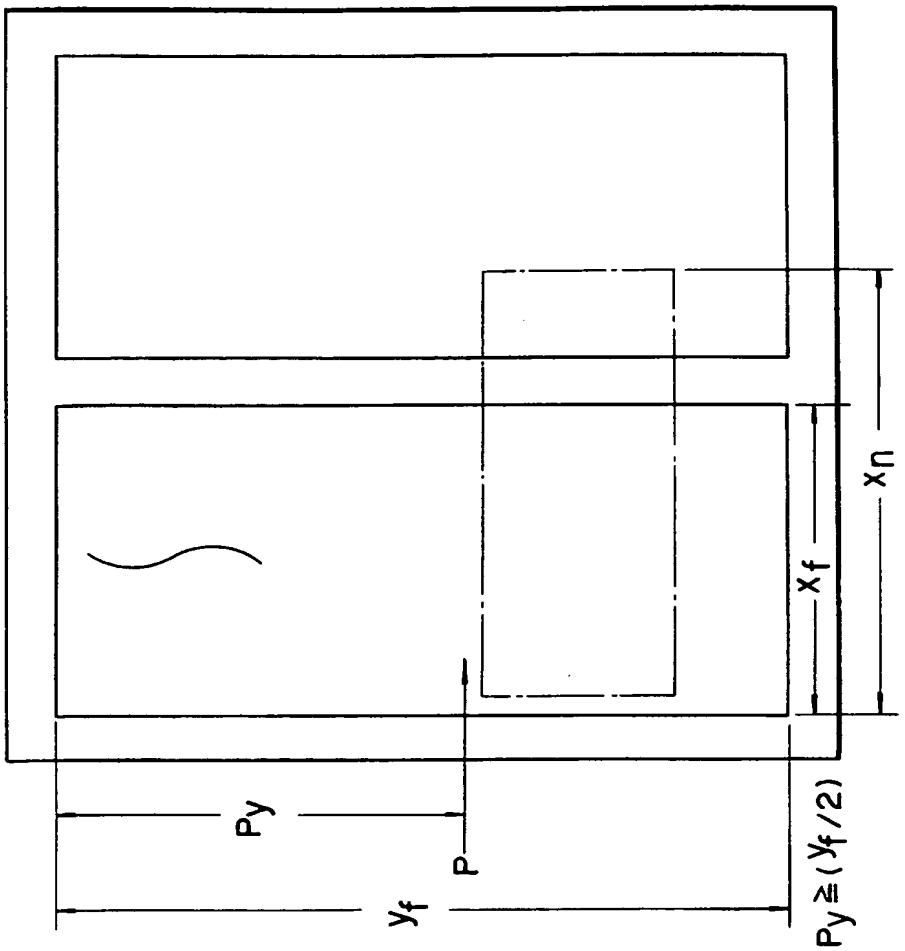


FIG. 7C

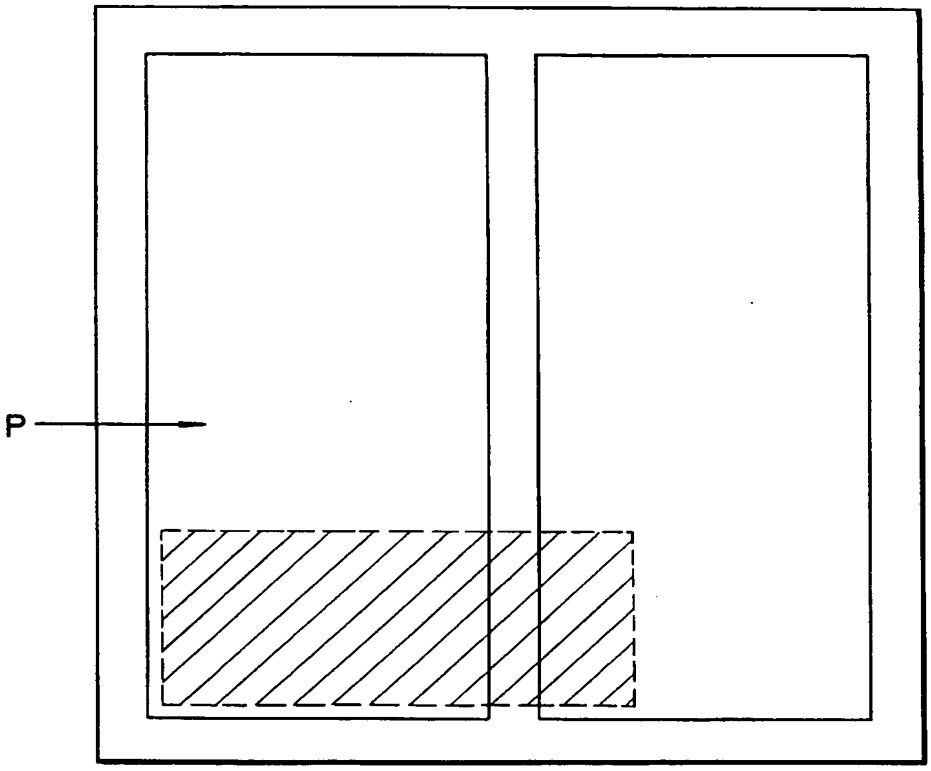


FIG. 7D

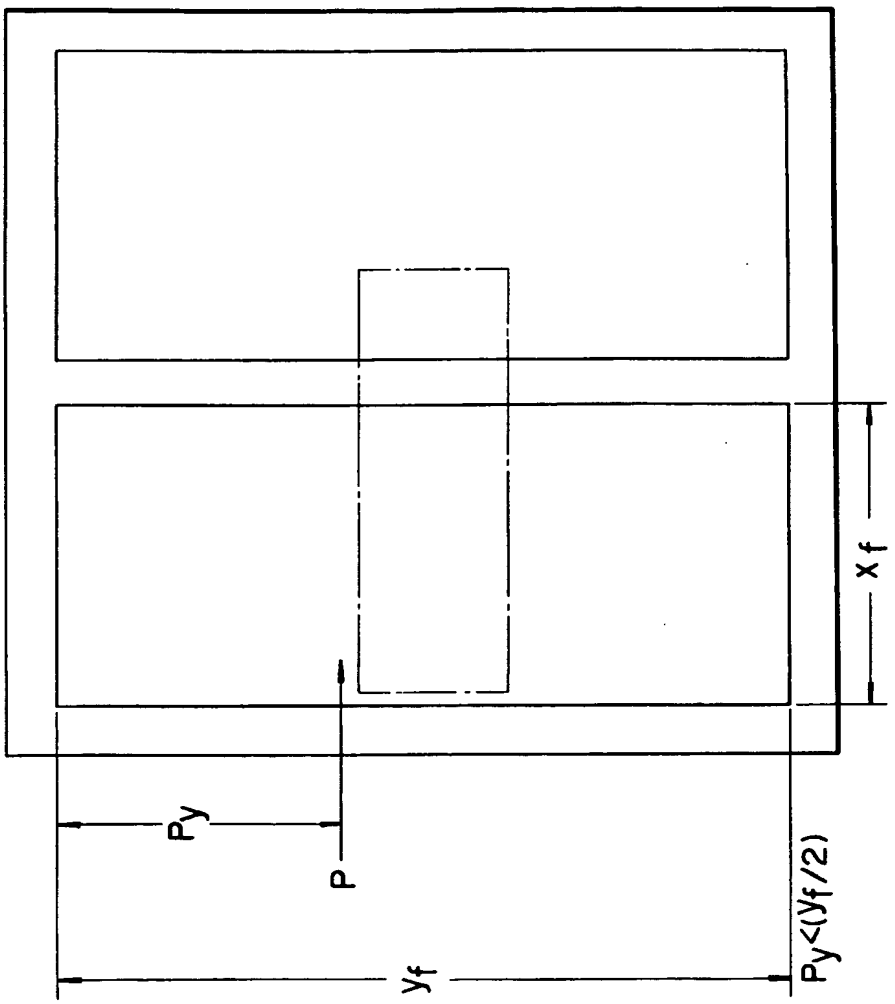


FIG. 7E

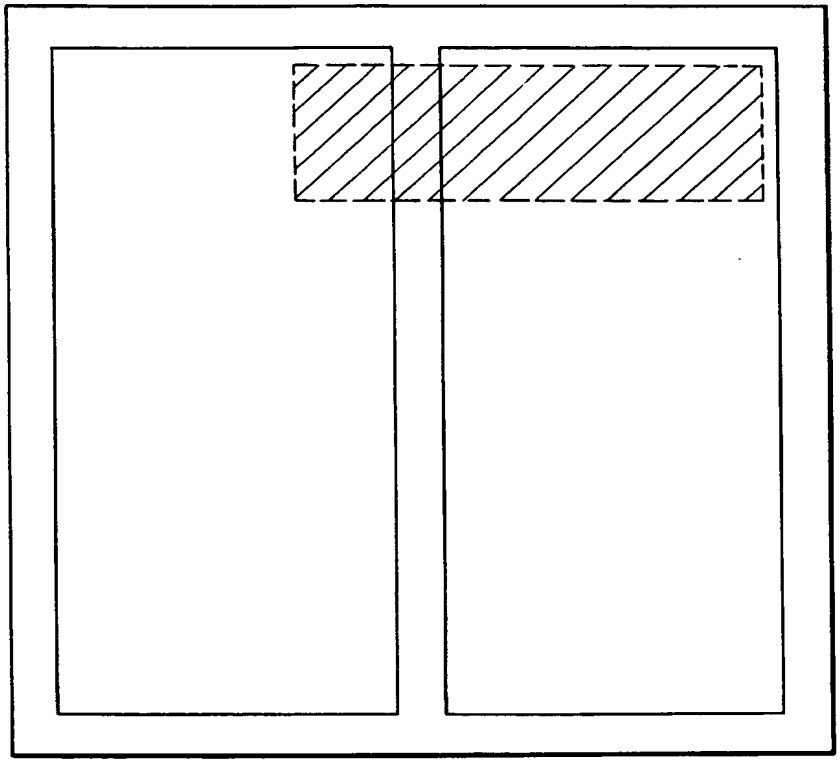

FIG. 7F

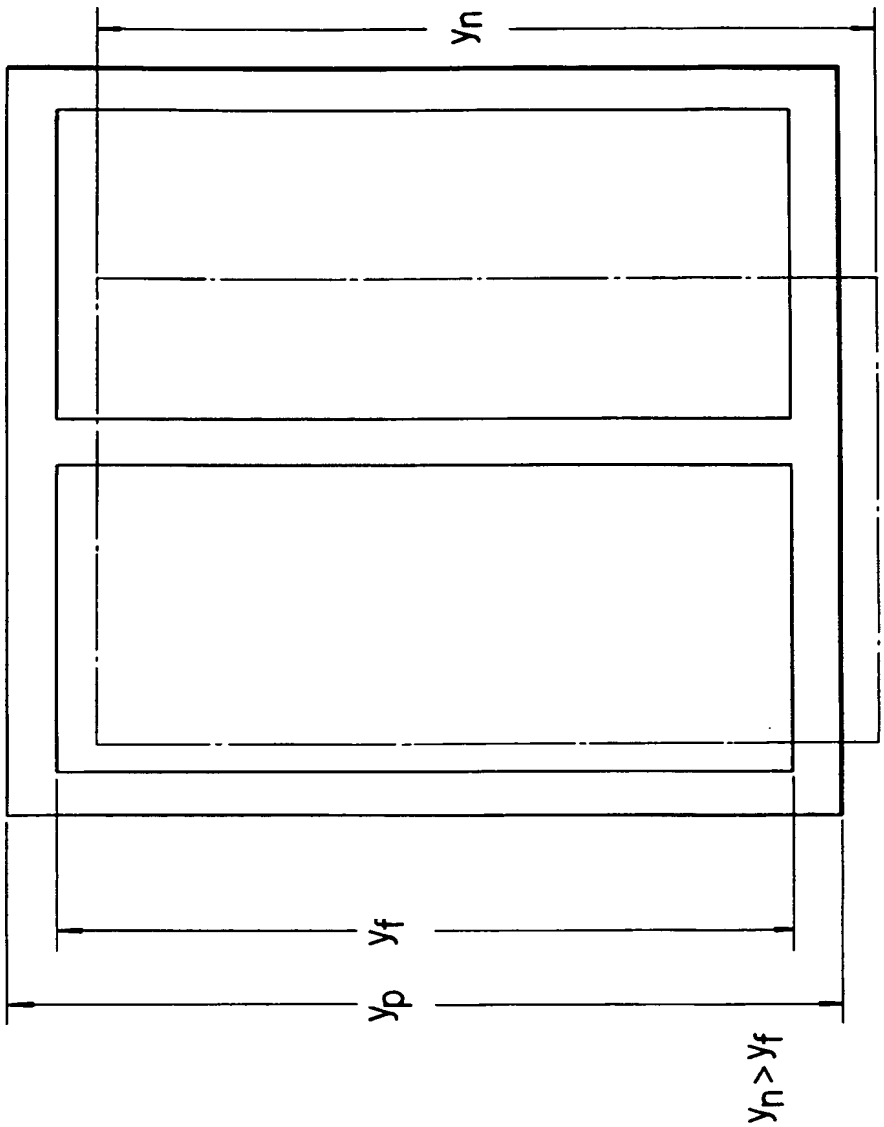


FIG. 7G

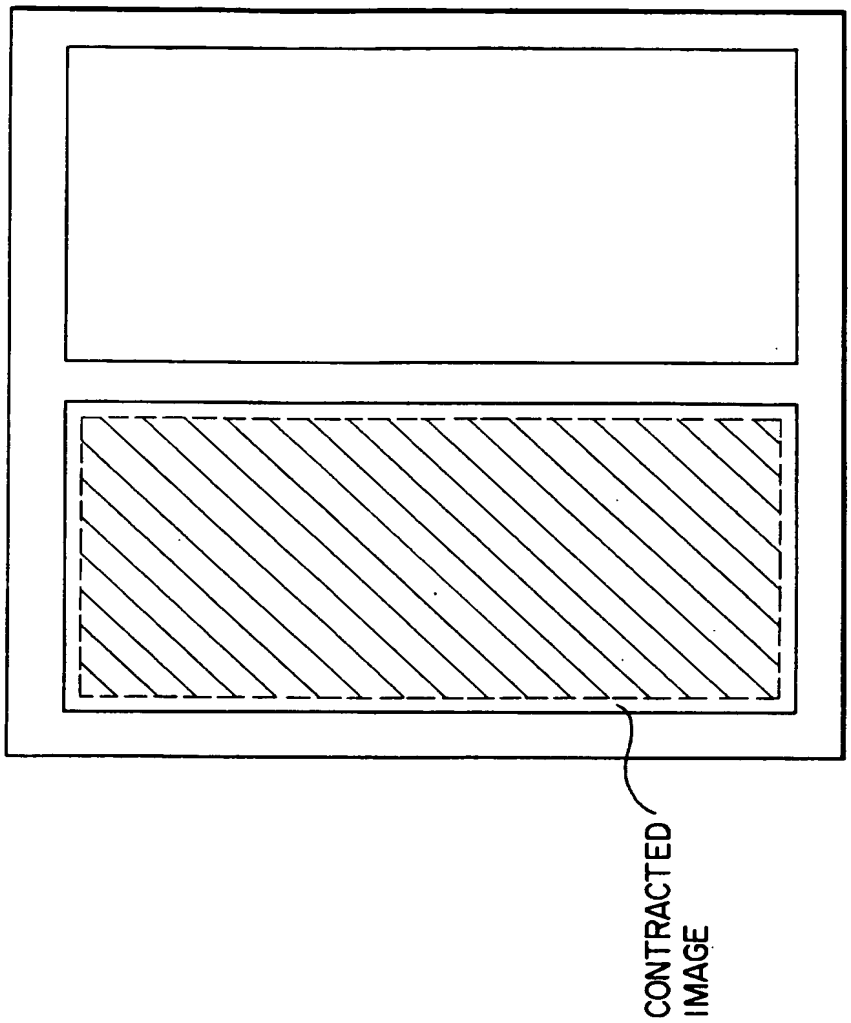


FIG. 7H